International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Feature Extraction and Analysis of SIFT features for ELA of Authentic and Forged Images

Rupali M. Bora[1*], Mahesh R. Sanghavi[2]

Submitted: 06/12/2023 **Revised**: 12/01/2024 **Accepted**: 30/01/2024

Abstract: With the advancements of technology in current era, everyone faces a challenge to identify digitally manipulated images. It is not easy to discriminate the original and forged images. For digital image tampering, image splicing and copy-move forgeries are very much well-known and common techniques. Image forgery is detected and spotted based on feature descriptor of an image. It is a concise and important local descriptor which is to be applied to grasp hierarchical representations from the input images. The significant correlation among nearby pixels has been identified by deep learning-based methods. It prefers locally grouped networks rather than one-to-one networks among all pixels. In the conducted research, the primary objective was to discern the authenticity of images using an integrated approach involving Error Level Analysis (ELA), Scale-Invariant Feature Transform (SIFT) features, and specific considerations for image types. These are Authentic, Copy-Move, and Spliced. The images under investigation were exclusively of JPEG format with size of 384x256 pixels. The Average SIFT feature values for the Authentic images consistently surpassed those of both Copy-Move and Spliced images. This discrepancy in feature values across the three categories, considering the standardized image format, presented a distinct opportunity for classification.

Keywords: *Copy-move, Error Level Analysis (ELA), Image forgery, Scale Invariant Feature Transform (SIFT), Spliced images*

1. Introduction

It is very easy to create fake and manipulated images because there are many easy image manipulation software and tools. With the advancements of technology in current era, the misuse of data in the form of images is in boom everywhere. All of us face this thought-provoking task of identifying such manipulated images and separating the real, called as pristine images, from the forged images. The most common technique used for digital image manipulation is splicing. In this a selected area is taken from an image and it is put on same or another one. It is reliant on proofs and evidences that are caused due to manipulation of images. Some general evidences are object inconsistency, lighting conditions and edge discontinuity. If same area of an image is used then it is a copy-move forgery. It may be done to create illusion of multiple objects of same type which already exist in the image. If a correlation value of entire image is computed, then it found that this value of these two regions of this manipulated image will be comparatively greater considering other areas of that image. Correct recognition of the copies in the images is the main aim of copy-move forgery. Various distance measures can be used for comparison of attributes extracted from image features. It is very difficult to detect and perceive tampering visually. An effective solution is required for tampering detection problem. There are many application Areas like: Surveillance systems, Intelligence services, medical imaging, Journalism, Forensic study, and Criminal search. A sample splicing forgery is shown in Fig. 1a, 1b below, where the animal zebra is copied to new background. A sample copy-move forgery is shown in Fig. 2a, 1b below, where the girl on stairs is duplicated on the same image. Most research work was proposed on image splicing, as it's easy to detect image inconsistency in illumination direction, contrast and noise which causes to detect tampering traces using deep learning based models.

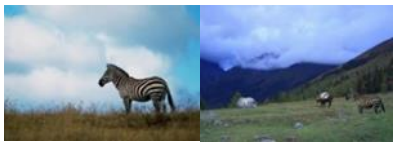

Fig. 1. a)Authentic image b)Image splicing forgery of 1a [1]

Fig. 2. a) Original image [1] b) Copy-move forgery of 2a [1]

Many types of methods and techniques are proposed for detection of the image forgeries. Generally, the image forgery identification techniques detect it by finding the

[1]*Research Scholar, Department of Computer Engineering, MET Institute of Engineering, Nashik affiliated to Savitribai Phule Pune University, India*
ORCID ID: 0000-0001-9916-597X
[2] *Professor, Department of Computer Engineering, SNJB's L.S.K.B.J. College of Engineering, Chandwad, Nashik, India*
ORCID ID: 0000-0001-7482-7455

dissimilarity of various attributes in image. There are some properties of images like contrast, illumination, shadow, compression, sensor noise which are important for this task. For various computer vision tasks Deep learning-based models are very much popular recently. Many images related tasks can be done using DL based models like object recognition, segmentation, and classification of images. There are two major reasons for Deep learning-based models' success in computer vision. Firstly, significant correlation between adjacent pixels is typically taken advantage. So, the grouping of networks locally than over one-to-one networks among all pixels is more preferred. Second, through a convolution operation each output feature map is formed by weights sharing. The typical methods are dependent on engineered attributes for detection of exact forgery. While the DL based models are based on training of existing images and identification of newly and different types of manipulations. These advantages of deep learning models help to find the existence of manipulation and forgery in an image. The artifacts found in the image are used to train the model.

2. Related Work

In the work of Sudiatmika et al. (2019) [2], for detecting the compression ratio is computed for original image and fake image using Error Level Analysis. VGG-16 architecture was modelled for identification of image manipulation through Error Level Analysis. The experimentation was carried out on CASIA_v2 dataset with accuracy of 92.2% for training and 88.46% for validation in 100 epochs. In the work of Liu, Y. H. (2018, September) [3], complete recognition process of CNN was presented. Due to local connection and weight sharing, CNN scan and extract features at very low computation cost. Due to pooling robustness of the network get enhanced. In the work of Mahale et al. (2017) [4], the method for image inconsistency detection was based on Binary Pattern. Experimentation was carried out on COMOFOD dataset and the efficiency was calculated using True positive Rate (TPR) and False Positive Rate (FPR). For different block sizes 2x2, 4x4, 8x8, 16x16 TPR values were in the range of 0.0142 to 0.08 and FPR values were in the range of 0.0995 to 0.0997 for these block sizes. The best results were shown for LBP of 2x2 block size as compared other sizes. In the work of Mareen et. al (2022, August) [5], uniqueness was presented by use of compression fingerprint which is representation of compression history. Training was done for only pristine data. Different forgeries are exposed by detection and localization of inconsistencies in the compression fingerprint. A fusion of Comprint with Noiseprint was implemented to confirm highest efficiency. In the work of Chakraborty et. al (2022) [6], authentic and tampered images can be differentiated using a dual-branch CNN. Preprocessing was done at beginning for evaluating the Error Level Analysis (ELA) of images. It then fed it to one branch of the network. Thirty high-pass filters were used for computation of noise residuals. The highest accuracy was achieved with smaller number of parameters which has minimum time and space complexity. In the work of Rhee, K. H. (2021) [7], ground truth image were generated for Copy-Move images. From Copy-Move forged images classification and semantic segmentation tasks can be achieved. The forgery patches were generated by applying various transformations like scaling, rotating, and blurring, etc. The method based on CNN helped to improve on accuracy and F1-score for image classification and semantic segmentation. In the work of Azhan, N. A. N., Ikuesan, R. A., Razak, S. A., & Kebande, V. R. (2022) [8], JPEG block signature was identified using error level analysis was proposed. 8×8 blocks are processed for different JPEG compression stages which were based on ELA values. These values had range of 0 to 3.0. In the work of Lowe, D. G. (2004) [9], distinctive SIFT keypoints were described which allows correct match for a keypoint to be designated from many keypoints. The keypoints found to be invariant and stable considering rotation and scaling. Also, they were robust across a considerable range of affine distortion, addition of noise, and changes in illumination. More number of keypoints extracted from images allows extracting small objects as well. But if the keypoints were detected over different scales then local features allows matching of small and highly concentrated objects. The methods for object recognition using keypoints were also discussed. In the work of Alberry et. al (2018) [10], optimization of Fuzzy C-Means technique was carried out for clustering SIFT key points with minimum time complexity. It also detects existence of geometric transformations and several copy-move forged images. A new dataset was generated for Copy-move forgery detection that includes more manipulations in images by professionals. Optimization of multiple clustering algorithms was done by matrix optimization. In the work of Liu, X., Liu, Y., Chen, J., & Liu, X. (2022) [11], for forgery detection, a Progressive Spatio-Channel Correlation Network (PSCC-Net) was developed. It perform detection along with localization of image manipulations. Two-path procedure is used here. Local and global features are extracted in the top-down path and image manipulation is detected and their manipulation mask is also detected at four scales in bottom-up path. Spatio-Channel Correlation Module (SCCM) was used in bottom-up path, which collects two types of correlations viz. spatial and channel-wise. This enables the model to manage with a extensive variety of manipulation attacks. Experimentation was conducted on synthetic dataset generated using MS COCO image dataset. In the work of Ali et. al (2022) [12], a robust system was developed for twice image compression context. For the original and

recompressed image, the variation is computed which is used for training purposes. This CNN-based model was lightweight and trivial. The performance of this model exhibits its great efficiency. The experimental results give validation accuracy of 92.23% on CASIA_v2 dataset. In the work of Kwon, M. J., Yu, I. J., Nam, S. H., & Lee, H. K. (2021) [13], CAT-Net, a CNN with RGB and DCT streams was developed. For RGB and DCT domains the forensic features of compression artifacts were learned together by this model. For handling various shapes and sizes of spliced object multiple resolutions of each stream are considered. Pretraining of the DCT stream allows to make use of its artifacts for double JPEG detection. Experimentation was conducted on CASIA_v2, Spliced COCO, Fantastic Reality, IMD2020 image datasets. In the work of Bappy et. al (2019) [14], high-confidence manipulation localization architecture was modelled. This architecture utilizes multiple frameworks like resampling features, LSTM and network of encoder-decoder. For image tamper localization the mapping from low-resolution features to pixel-wise predictions was used. Along with this, final layer predicts the mask. The pixel level image manipulations are localized with high precision. For learning correspondence of encoded features to binary mask, a decoder network is implemented. A new synthesized dataset of images was introduced. The detailed experiments showed an efficient separation of different forgeries like image splicing, copy-move and object-removal. The performance was tested on IEEE Forensics, NIST16, COVERAGE image datasets. In the work of Yancey, R. E. (2019) [15], Faster RCNN network was based on Multi-stream. The input of the element-wise sum of ELA along with Block Artifact Grid (BAG) error provide high accuracy. Performance was evaluated on CASIA_v2, CoMoFoD, COVERAGE datasets with maximum accuracy of 82%.

3. Methodology

3.1 Steps for feature extraction and ELA features analysis

The steps carried out to analyse the image features based on ELA are as follows:

i. Read a JPEG input image
ii. Find ELA for various Q factors (97, 94, 91, 88, 85, 82)
iii. Find SIFT features for each ELA output image for each Q factor
iv. Generate csv files for SIFT features
v. Generate Count of SIFT features
vi. Generate average for each Q factor for all images

The overall process is as shown in Fig. 3. in the block diagram for feature extraction and analysis.

3.2 Error Level Analysis (ELA)

In the processing of images for forensic analysis, ELA has been extensively used. From the discrepancy between original images its compressed image, error pattern could be inferred. This error is computed based on 8×8 blocks. In ELA, an image is resaved for a particular level of compression quality. The observations of compression levels' differences were analyzed. The computation of ELA levels for each block of JPEG (i,j) is given in equation (1) [8].

ELA level = Resaved Image – Recompressed Image

$$I_{75}(x,y)-I_{95'}(x,y)=ELA_1$$

$$I_{75'}(x,y)-I_{95''}(x,y)=ELA_2$$

$$I_{75''}(x,y)-I_{95'''}(x,y)=ELA_3$$

$$\ldots\ldots$$

$$I_{75n}(x,y)-I_{95n}(x,y)=ELA_n \quad (1)$$

Here 'I' means a JPEG image. Like I_{75} signifies a JPEG resaved with 75% quality factor, and I_{95} signifies a JPEG recompressed with 95% quality factor. Here (') dash is used to indicate recompression 1,2,…n times. The compression error denoted by ELA1,2,...,n get decreased as JPEG is resaved for multiple times. Local minima is be achieved slowly due to resaving of images for 1,2,…n times. Due to this darkness will be seen for each JPEG block 8x8.

3.3 Scale Invariant Feature Transform (SIFT)

This is a well-known image feature detection and description technique. Distinctive keypoints in an image are detected using SIFT. These key points are robust considering rotation, scale and affine transformations. Based on their local intensity extrema, these key points are identified. The complete process can be carried out 4 steps[9]:

i. Constructing a Scale Space: to detect the most discrete features in a given input image without considering noise and to guarantee that features are scale-independent, scale-space is constructed. It is a group of images with different scaling for an image. Noise reduction is done using Gaussian blur and finding Difference of this Gaussian across scales.

ii. Keypoint Localisation: For recognizing the appropriate features i.e. keypoints, the local maxima and minim is found and low contrast keypoints are removed to generate scale-invariant keypoints.

iii. Orientation Assignment: For confirmation that the keypoints are rotation-invariant, magnitude and direction are calculated. Then a histogram is plotted for magnitude and orientation.

iv. Keypoint Descriptor: From the keypoint, their neighboring pixels along with its magnitude and direction are used to produce a unique pattern for this keypoint which is called as 'descriptor'.

These steps generate magnitude and direction for descriptor which are the SIFT features.

4. Results and Discussion

4.1 Experimentation

From the CASIA_v2 dataset [1] JPEG images of size of 384x256 pixels were selected. Here, we have taken 380 authentic, 124 copy-move and 93 spliced images.

Fig. 3. Block diagram for feature extraction and analysis

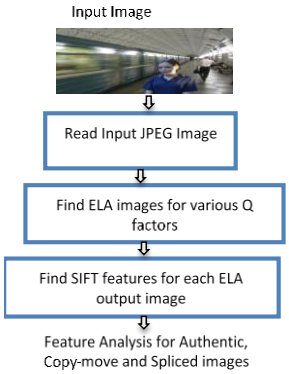

Fig. 4. SIFT features for ELA images of different Q factors

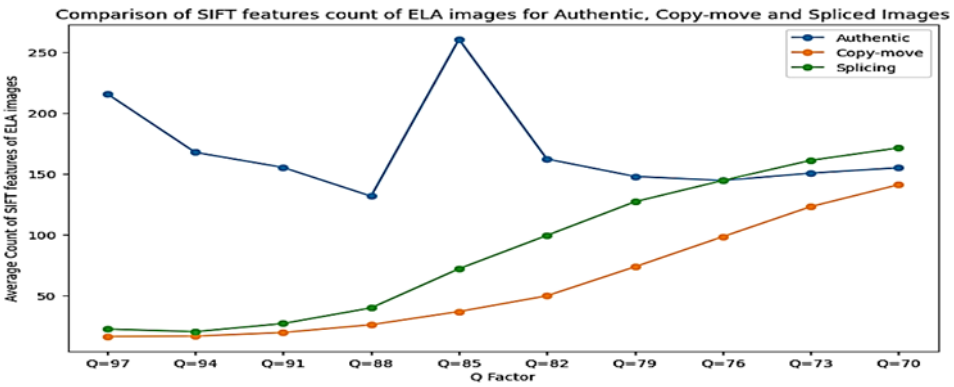

Fig. 5 Comparison of SIFT features for ELA images of different Q factors

In the conducted research, the primary objective was to discern the authenticity of images using an integrated approach involving ELA, SIFT features, and specific considerations for image types. Three distinct image categories were considered: Authentic, Copy-Move, and Spliced. The images under investigation were exclusively of JPEG format with a standardized size of 384x256 pixels. The procedure involved the extraction of ELA images at varying quality levels (97, 94, 91, 88, 85, 82, 79, 76, 73, 70) for each image category, followed by the calculation of SIFT features for each ELA image as shown in Fig. 4. The values are computed are as shown in Table 1. Upon thorough examination of the obtained results, a notable trend emerged. The Average SIFT feature values for the authentic images consistently surpassed those of both Copy-Move and Spliced images as is shown in Fig. 5.

4.2 Result Analysis

This discrepancy in feature values across the three categories, considering the standardized image format, presented a distinct opportunity for classification. To capitalize on this observation, a threshold value of 100 was established. Any image surpassing this threshold in terms of the average SIFT feature value across all quality levels was categorized as Authenticated, while those falling below were deemed Unauthenticated. This approach was substantiated by a comprehensive analysis of the data. Notably, the Authenticated images exhibited higher SIFT feature values, signifying a greater complexity in terms of distinctive keypoints.

Table 1. SIFT features Counts for ELA images of different Q factors

Q factor for ELA	97	94	91	88	85	82	79	76	73	70
Authentic Images	216	168	155	132	261	162	148	145	150	155
Copy-move Images	16	17	20	26	37	50	74	98	123	141
Spliced Images	22	20	27	40	72	99	127	145	161	171

Conversely, Copy-Move and Spliced images, being manipulated or composite in nature, exhibited comparatively lower average SIFT feature values. The methodological framework outlined above, tailored for JPEG images of dimensions 384x256 pixels, holds promise for practical implementation. It provides a nuanced understanding of image integrity with a focus on both global and local characteristics, specific to the standardized image type under consideration. The established threshold of 100 serves as a reliable discriminator between authentic and manipulated images, offering a valuable contribution to image forensics.

5. Conclusion

For analysis of image features three categories of images have been used viz. Authentic, Copy-Move, and Spliced, with JPEG format and image dimension of 384x256 from CASIA_v2 dataset. The ELA images are generated for 10 different quality levels (97, 94, 91, 88, 85, 82, 79, 76, 73, 70). The SIFT features for each ELA image are extracted for analysis. The experimentation shows that the average SIFT feature values for the Authentic images consistently surpassed those of both Copy-Move and Spliced images for all 10 quality levels. In future, this analysis will be used to further classify the images into three classes as authentic, copy-move and spliced. This research contributes to the burgeoning field of image forensics, with potential applications in diverse domains such as security, law enforcement, and digital media authentication.

Acknowledgements

We thank our colleagues from K. K. Wagh Institute of Engineering Education and Research who provided insight and expertise that greatly assisted the research. And special thanks to Shrikant S. Pawar, a student from class BEIT for his contribution in the work of Software and Visualization.

Author contributions

Rupali M. Bora: Conceptualization, Methodology, Software, Writing-Original draft preparation **Mahesh R. Sanghavi:** Investigation, Writing-Reviewing Data curation, Validation

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Dong, J., Wang, W., & Tan, T. (2013, July). Casia image tampering detection evaluation database. In *2013 IEEE China summit and international conference on signal and information processing* (pp. 422-426). IEEE.

[2] Sudiatmika, I. B. K., Rahman, F., Trisno, T., & Suyoto, S. (2019). Image forgery detection using error level analysis and deep learning. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *17*(2), 653-659.

[3] Liu, Y. H. (2018, September). Feature extraction and image recognition with convolutional neural networks. In *Journal of Physics: Conference Series* (Vol. 1087, p. 062032). IOP Publishing.

[4] Mahale, V. H., Ali, M. M., Yannawar, P. L., & Gaikwad, A. T. (2017). Image inconsistency detection using local binary pattern (LBP). *Procedia computer science*, *115*, 501-508.

[5] Mareen, H., Vanden Bussche, D., Guillaro, F., Cozzolino, D., Van Wallendael, G., Lambert, P., & Verdoliva, L. (2022, August). Comprint: Image forgery detection and localization using compression fingerprints. In *International Conference on Pattern Recognition* (pp. 281-299). Cham: Springer Nature Switzerland.

[6] Chakraborty, S., Chatterjee, K., & Dey, P. (2022). Detection of Image Tampering Using Deep Learning, Error Levels & Noise Residuals..

[7] Rhee, K. H. (2021). Generation of novelty ground truth image using image classification and semantic segmentation for copy-move forgery detection. *IEEE Access*, *10*, 2783-2796.

[8] Azhan, N. A. N., Ikuesan, R. A., Razak, S. A., & Kebande, V. R. (2022). Error Level Analysis Technique for Identifying JPEG Block Unique Signature for Digital Forensic Analysis. *Electronics*, *11*(9), 1468.

[9] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, *60*, 91-110.

[10] Alberry, H. A., Hegazy, A. A., & Salama, G. I. (2018). A fast SIFT based method for copy move forgery detection. *Future Computing and Informatics Journal*, *3*(2), 159-165.

[11] Liu, X., Liu, Y., Chen, J., & Liu, X. (2022). PSCC-Net: Progressive spatio-channel correlation network for image manipulation detection and localization. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(11), 7505-7517.

[12] Ali, S. S., Ganapathi, I. I., Vu, N. S., Ali, S. D., Saxena, N., & Werghi, N. (2022). Image Forgery Detection Using Deeplearning by Recompressing Images. Electronics 2022, 11, 403.

[13] Kwon, M. J., Yu, I. J., Nam, S. H., & Lee, H. K. (2021). CAT-Net: Compression artifact tracing network for detection and localization of image splicing. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 375-384).

[14] Bappy, J. H., Simons, C., Nataraj, L., Manjunath, B. S., & Roy-Chowdhury, A. K. (2019). Hybrid lstm and encoder–decoder architecture for detection of image forgeries. *IEEE Transactions on Image Processing*, *28*(7), 3286-3300.

[15] Yancey, R. E. (2019). Deep localization of mixed image tampering techniques. *arXiv preprint arXiv:1904.08484*.